

### SE, SENIOR ENGINEER

4110 · Spalding, Edgar P.  
Noh, Bos;

## 1120 > MDR Like ABC Transporter Gene From Plants

130-13238 00061

«140» herewith  
«141» not yet assigned

43502 PCT US99/22363

• 151 • 1399 09 24

<150> US 60/101,814

<151> 1998-09-25

<160> 14

<170> Fast SEQ for Windows Version 4.0

62105-1

<211> - 40.51

#### $\leq 15 \text{ } \mu\text{m}$ DNA

<211> *Arabidopsis thaliana*

220

<221> misc feature

222 (94) . . . 101

<223> Translation start codon

<221> misc feature

<222> (3932) . . . (0)

◀ 223 Stop codon

— 14 —



Net Ser Glu Thr Asn Thr Thr Asp Ala Lys Thr Val Pro Ala Glu Ala  
 1 5 10 15  
 Glu Lys Lys Lys Glu Cln Ser Leu Pro Phe Phe Lys Leu Phe Ser Phe  
 20 25 30  
 Ala Asp Lys Phe Asp Tyr Leu Leu Met Phe Val Gly Ser Leu Gly Ala  
 35 40 45  
 Ile Val His Gly Ser Ser Met Pro Val Phe Phe Leu Leu Phe Gly Cln  
 50 55 60  
 Met Val Asn Gly Phe Gly Lys Asn Cln Met Asp Leu His Cln Met Val  
 65 70 75 80  
 His Glu Val Ser Arg Tyr Ser Leu Tyr Phe Val Tyr Leu Gly Leu Val  
 85 90 95  
 Val Cys Phe Ser Ser Tyr Ala Glu Ile Ala Cys Trp Met Tyr Ser Gly  
 100 105 110  
 Glu Arg Gln Val Ala Ala Leu Arg Lys Lys Tyr Leu Glu Ala Val Leu  
 115 120 125  
 Lys Glu Asp Val Gly Phe Phe Asp Thr Asp Ala Arg Thr Gly Asp Ile  
 130 135 140  
 Val Phe Ser Val Ser Thr Asp Thr Leu Leu Val Gln Asp Ala Ile Ser  
 145 150 155 160  
 Glu Lys Val Gly Asn Phe Ile His Tyr Leu Ser Thr Phe Leu Ala Gly  
 165 170 175  
 Leu Val Val Gly Phe Val Ser Ala Trp Lys Leu Ala Leu Leu Ser Val  
 180 185 190  
 Ala Val Ile Pro Gly Ile Ala Phe Ala Gly Gly Leu Tyr Ala Tyr Thr  
 195 200 205  
 Leu Thr Gly Ile Thr Ser Lys Ser Arg Glu Ser Tyr Ala Asn Ala Gly  
 210 215 220  
 Val Ile Ala Glu Gln Ala Ile Ala Gln Val Arg Thr Val Tyr Ser Tyr  
 225 230 235 240  
 Val Gly Glu Ser Lys Ala Leu Asn Ala Tyr Ser Asp Ala Ile Gln Tyr  
 245 250 255  
 Thr Leu Lys Leu Gly Tyr Lys Ala Gly Met Ala Lys Gly Leu Gly Leu  
 260 265 270  
 Gly Cys Thr Tyr Gly Ile Ala Cys Met Ser Trp Ala Leu Val Phe Trp  
 275 280 285  
 Tyr Ala Gly Val Phe Ile Arg Asn Gly Gln Thr Asp Gly Gly Lys Ala  
 290 295 300  
 Ile Thr Ala Ile Ile Asn Ala Ile Ile Ile Gly Met Ser Leu Gly Cln  
 310 315 320  
 Ser Ile Asp Asn Leu Ile Asp Asn Ile Ile Tyr Gly Ile Ala Ala Ile Tyr  
 330 335 340  
 Ile Val Met Ile Ile Ile Asn Cln Arg Phe Val Ile Ile Cln Asp Phe  
 340 345 350  
 Leu Asp Gly Lys Cys Leu Asp Cln Val His Gly Asn Ile Glu Phe Lys  
 355 360 365  
 Asp Val Thr Phe Ser Tyr Pro Ser Arg Pro Asp Val Met Ile Phe Arg  
 370 375 380  
 Asn Phe Asp Ile Ile Ile Asp Ser Ile Ile Val Val Ala Ile Thr Ile  
 390 395 400

450                  455                  460  
Lys Pro Asp Ala Thr Met Ile Glu Val Glu Ala Ala Ala Ser Ala Ala  
465                  470                  475                  480  
Asn Ala His Ser Phe Ile Thr Leu Leu Pro Lys Gly Tyr Asp Thr Gln  
485                  490                  495  
Val Gly Glu Arg Gly Val Gln Leu Ser Gly Gly Gln Lys Gln Arg Ile  
500                  505                  510  
Ala Ile Ala Arg Ala Met Leu Lys Asp Pro Lys Ile Leu Leu Leu Asp  
515                  520                  525  
Glu Ala Thr Ser Ala Leu Asp Ala Ser Ser Glu Ser Ile Val Gln Glu  
530                  535                  540  
Ala Leu Asp Arg Val Met Val Gly Arg Thr Thr Val Val Val Ala His  
545                  550                  555                  560  
Arg Leu Cys Thr Ile Arg Asn Val Asp Ser Ile Ala Val Ile Gln Gln  
565                  570                  575  
Gly Gln Val Val Glu Thr Gly Thr His Glu Glu Leu Ile Ala Lys Ser  
580                  585                  590  
Gly Ala Tyr Ala Ser Leu Ile Arg Phe Gln Glu Met Val Gly Thr Arg  
595                  600                  605  
Asp Phe Ser Asn Pro Ser Thr Arg Arg Thr Arg Ser Thr Arg Leu Ser  
610                  615                  620  
His Ser Leu Ser Thr Lys Ser Leu Ser Ile Arg Ser Gly Ser Leu Arg  
625                  630                  635                  640  
Asn Leu Ser Tyr Ser Tyr Ser Thr Gly Ala Asp Gly Arg Ile Glu Met  
645                  650                  655  
Ile Ser Asn Ala Glu Thr Asp Arg Lys Thr Arg Ala Pro Glu Asn Tyr  
660                  665                  670  
Phe Tyr Arg Leu Leu Lys Leu Asn Ser Pro Glu Trp Pro Tyr Ser Ile  
675                  680                  685  
Met Gly Ala Val Gly Ser Ile Leu Ser Gly Phe Ile Gly Pro Thr Phe  
690                  695                  700  
Ala Ile Val Met Ser Asn Met Ile Glu Val Phe Tyr Tyr Thr Asp Tyr  
705                  710                  715                  720  
Asp Ser Met Glu Arg Lys Thr Lys Glu Tyr Val Phe Ile Tyr Ile Gly  
725                  730                  735  
Ala Gly Leu Tyr Ala Val Gly Ala Tyr Leu Ile Gln His Tyr Phe Phe  
740                  745                  750  
Ser Ile Met Gly Glu Asn Leu Thr Thr Arg Val Arg Arg Met Met Leu  
755                  760                  765  
Ser Ala Ile Leu Arg Asn Glu Val Gly Ile Phe Asp His Asp Ile His  
770                  775                  780  
Asn Ile Leu Leu Ile Ala Asn Ile Ala Asn Ile Asn Ile Asn Ile Asn  
785                  790                  795                  800  
Lys Ser Ala Ile Ala Asn Asn Arg Ile Ser Ile Ile Leu Glu Asn Met Ile  
805                  810                  815  
Ser Leu Leu Thr Ser Phe Ile Val Ala Phe Ile Val Glu Trp Arg Val  
820                  825                  830  
Ser Leu Leu Leu Ile Asp Thr Ile Pro Leu Ile Ile Leu Ala Asn Phe  
835                  840                  845  
Ala Thr Ile Leu Phe Ile Ile Asp Ile Ala Thr Ala Ile Ile Ala

Gln Thr Ser Gly Phe Leu Phe Gly Leu Ser Gln Leu Ala Leu Tyr Gly  
 915 920 925  
 Ser Glu Ala Leu Ile Leu Trp Tyr Gly Ala His Leu Val Ser Lys Gly  
 930 935 940  
 Val Ser Thr Phe Ser Lys Val Ile Lys Val Phe Val Val Leu Val Ile  
 945 950 955 960  
 Thr Ala Asn Ser Val Ala Glu Thr Val Ser Leu Ala Pro Glu Ile Ile  
 965 970 975  
 Arg Gly Gly Glu Ala Val Gly Ser Val Phe Ser Val Leu Asp Arg Gln  
 980 985 990  
 Thr Arg Ile Asp Pro Asp Asp Ala Asp Ala Asp Pro Val Glu Thr Ile  
 995 1000 1005  
 Arg Gly Asp Ile Glu Phe Arg His Val Asp Phe Ala Tyr Pro Ser Arg  
 1010 1015 1020  
 Pro Asp Val Met Val Phe Arg Asp Phe Asn Leu Arg Ile Arg Ala Gly  
 1025 1030 1035 1040  
 His Ser Gln Ala Leu Val Gly Ala Ser Gly Ser Gly Lys Ser Ser Val  
 1045 1050 1055  
 Ile Ala Met Ile Glu Arg Phe Tyr Asp Leu Leu Ala Gly Lys Val Met  
 1060 1065 1070  
 Ile Asp Gly Lys Asp Ile Arg Arg Leu Asn Leu Lys Ser Leu Arg Leu  
 1075 1080 1085  
 Lys Ile Gly Leu Val Gln Gln Glu Pro Ala Leu Phe Ala Ala Thr Ile  
 1090 1095 1100  
 Phe Asp Asn Ile Ala Tyr Gly Lys Asp Gly Ala Thr Glu Ser Glu Val  
 1105 1110 1115 1120  
 Ile Asp Ala Ala Arg Ala Ala Asn Ala His Gly Phe Ile Ser Gly Leu  
 1125 1130 1135  
 Pro Glu Gly Tyr Lys Thr Pro Val Gly Glu Arg Gly Val Gln Leu Ser  
 1140 1145 1150  
 Gly Gly Gln Lys Arg Ile Ala Ile Ala Arg Ala Val Leu Lys Asn  
 1155 1160 1165  
 Pro Thr Val Leu Leu Leu Asp Glu Ala Thr Ser Ala Leu Asp Ala Glu  
 1170 1175 1180  
 Ser Glu Cys Val Leu Gln Glu Ala Leu Glu Arg Leu Met Arg Gly Arg  
 1185 1190 1195 1200  
 Thr Thr Val Val Val Ala His Arg Leu Ser Thr Ile Arg Gly Val Asp  
 1205 1210 1215  
 Gly Ile Gly Val Ile Ile Asp Gly Arg Ile Val Gly Glu Ser His  
 1220 1225 1230 1235  
 Ser Val Leu Val Ser Asn Ile Ile Gly Ala Thr Ser Asp Ile Ile Ile  
 1240 1245 1250 1255  
 Leu Gln Thr His Arg Ile  
 1260

<210> 3  
 <211> 1260  
 <212> PPT  
 <213> Hemagglutinin

Phe Lys Leu Asn Asn Lys Ser Glu Lys Asp Lys Lys Glu Lys Lys Pro  
     20                       25                       30  
 Thr Val Ser Val Phe Ser Met Phe Arg Tyr Ser Asn Trp Leu Asp Lys  
     35                       40                       45  
 Leu Tyr Met Val Val Gly Thr Leu Ala Ala Ile Ile His Gly Ala Gly  
     50                       55                       60  
 Leu Pro Leu Met Met Leu Val Phe Gly Glu Met Thr Asp Ile Phe Ala  
     65                       70                       75                       80  
 Asn Ala Gly Asn Leu Glu Asp Leu Met Ser Asn Ile Thr Asn Arg Ser  
     85                       90                       95  
 Asp Ile Asn Asp Thr Gly Phe Phe Met Asn Leu Glu Glu Asp Met Thr  
     100                       105                       110  
 Arg Tyr Ala Tyr Tyr Ser Gly Ile Gly Ala Gly Val Leu Val Ala  
     115                       120                       125  
 Ala Tyr Ile Gin Val Ser Phe Trp Cys Leu Ala Ala Gly Arg Gln Ile  
     130                       135                       140  
 His Lys Ile Arg Lys Gln Phe Phe His Ala Ile Met Arg Gln Glu Ile  
     145                       150                       155                       160  
 Gly Trp Phe Asp Val His Asp Val Gly Glu Leu Asn Thr Arg Leu Thr  
     165                       170                       175  
 Asp Asp Val Ser Lys Ile Asn Glu Val Ile Gly Asp Lys Ile Gly Met  
     180                       185                       190  
 Phe Phe Gln Ser Met Ala Thr Phe Phe Thr Gly Phe Ile Val Gly Phe  
     195                       200                       205  
 Thr Arg Gly Trp Lys Leu Thr Leu Val Ile Leu Ala Ile Ser Pro Val  
     210                       215                       220  
 Leu Gly Leu Ser Ala Ala Val Trp Ala Lys Ile Leu Ser Ser Phe Thr  
     225                       230                       235                       240  
 Asp Lys Glu Leu Leu Ala Tyr Ala Lys Ala Gly Ala Val Ala Glu Glu  
     245                       250                       255  
 Val Leu Ala Ala Ile Arg Thr Val Ile Ala Phe Gly Gly Gln Lys Lys  
     260                       265                       270  
 Glu Leu Glu Arg Tyr Asn Lys Asn Leu Glu Glu Ala Lys Arg Ile Gly  
     275                       280                       285  
 Ile Lys Lys Ala Ile Thr Ala Asn Ile Ser Ile Gly Ala Ala Phe Leu  
     290                       295                       300  
 Leu Ile Tyr Ala Ser Tyr Ala Leu Ala Phe Trp Tyr Gly Thr Thr Leu  
     305                       310                       315                       320  
 Ile Leu Ser Gly Glu Tyr Ser Ile Gly Gin Val Leu Thr Val Phe Phe  
     325                       330                       335                       340  
 Ser Val Leu Ile Gly Ala Phe Phe Val Gly Ile Ala Ser Phe Ser Ile  
     345                       350                       355                       360  
 Ala Ala Phe Ala Asn Ala Asn Ala Ala Tyr Ile Ile Phe Lys Ile  
     365                       370                       375                       380  
 Ile Asp Asn Lys Pro Ser Ile Asp Ser Tyr Ser Lys Ser Gly His Lys  
     385                       390                       395                       400  
 Pro Asp Asn Ile Lys Gly Asn Leu Glu Phe Arg Asn Val His Phe Ser  
     405                       410                       415                       420  
 Ile Asp Asn Ile Lys Gly Asn Leu Glu Phe Arg Asn Val His Phe Ser  
     425                       430                       435                       440

465	470	475	480
Ala Thr Thr Ile Ala Glu Asn Ile Arg Tyr Gly Arg Glu Asn Val Thr			
485	490	495	
Met Asp Glu Ile Glu Lys Ala Val Lys Glu Ala Asn Ala Tyr Asp Phe			
500	505	510	
Ile Met Lys Leu Pro His Lys Phe Asp Thr Leu Val Gly Glu Arg Gly			
515	520	525	
Ala Gln Leu Ser Gly Gly Gln Lys Gln Arg Ile Ala Ile Ala Arg Ala			
530	535	540	
Leu Val Arg Asn Pro Lys Ile Leu Leu Asp Glu Ala Thr Ser Ala			
545	550	555	560
Leu Asp Thr Glu Ser Glu Ala Val Val Gln Val Ala Leu Asp Lys Ala			
565	570	575	
Arg Lys Gly Arg Thr Thr Ile Val Ile Ala His Arg Leu Ser Thr Val			
580	585	590	
Arg Asn Ala Asp Val Ile Ala Gly Phe Asp Asp Gly Val Ile Val Glu			
595	600	605	
Lys Gly Asn His Asp Glu Leu Met Lys Glu Lys Gly Ile Tyr Phe Lys			
610	615	620	
Leu Val Thr Met Gln Thr Ala Gly Asn Glu Val Glu Leu Glu Asn Ala			
625	630	635	640
Ala Asp Glu Ser Lys Ser Glu Ile Asp Ala Leu Glu Met Ser Ser Asn			
645	650	655	
Asp Ser Arg Ser Ser Leu Ile Arg Lys Arg Ser Thr Arg Arg Ser Val			
660	665	670	
Arg Gly Ser Gln Ala Gln Asp Arg Lys Leu Ser Thr Lys Glu Ala Leu			
675	680	685	
Asp Glu Ser Ile Pro Pro Val Ser Phe Trp Arg Ile Met Lys Leu Asn			
690	695	700	
Leu Thr Glu Trp Pro Tyr Phe Val Val Gly Val Phe Cys Ala Ile Ile			
705	710	715	720
Asn Gly Gly Leu Gln Pro Ala Phe Ala Ile Ile Phe Ser Lys Ile Ile			
725	730	735	
Gly Val Phe Thr Arg Ile Asp Asp Pro Glu Thr Lys Arg Gln Asn Ser			
740	745	750	
Asn Leu Phe Ser Leu Leu Phe Leu Ala Leu Gly Ile Ile Ser Phe Ile			
755	760	765	
Thr Phe Phe Leu Gln Gly Phe Thr Phe Gly Lys Ala Gly Glu Ile Leu			
770	775	780	
Arg Lys Arg Ile Arg Tyr Met Val Phe Asn Ile Met Leu Asp Ile Asp			
785	790	795	800
Leu Val Thr Ile Asp Asp Ile Asp Asn Ile Thr Ile Asn Leu Thr Val			
805	810	815	
Arg Leu Ala Asn Asp Ala Ala Gln Val Lys Gly Ala Ile Gly Ser Arg			
820	825	830	
Leu Ala Val Ile Thr Gln Asn Ile Ala Asn Leu Gly Thr Gly Ile Ile			
835	840	845	
Ile Leu Ile Ile Tyr Asp Trp Gln Leu Thr Leu Leu Leu Leu Ala Ile			
850	855	860	
Arg Leu Ile Ile Asp			

Arg Asn Ser Leu Arg Lys Ala His Ile Phe Gly Ile Thr Phe Ser Phe  
 930 935 940  
 Thr Gln Ala Met Met Tyr Phe Ser Tyr Ala Gly Cys Phe Arg Phe Gly  
 945 950 955 960  
 Ala Tyr Leu Val Ala His Lys Leu Met Ser Phe Glu Asp Val Leu Leu  
 965 970 975  
 Val Phe Ser Ala Val Val Phe Gly Ala Met Ala Val Gly Gln Val Ser  
 980 985 990  
 Ser Phe Ala Pro Asp Tyr Ala Lys Ile Ser Ala Ala His Ile  
 995 1000 1005  
 Ile Met Ile Ile Glu Lys Thr Pro Leu Ile Asp Ser Tyr Ser Thr Glu  
 1010 1015 1020  
 Gly Leu Met Pro Asn Thr Leu Glu Gly Asn Val Thr Phe Gly Glu Val  
 1025 1030 1035 1040  
 Val Phe Asn Tyr Pro Thr Arg Pro Asp Ile Pro Val Leu Gln Gly Leu  
 1045 1050 1055  
 Ser Leu Glu Val Lys Lys Gly Gln Thr Leu Ala Leu Val Gly Ser Ser  
 1060 1065 1070  
 Gly Cys Gly Lys Ser Thr Val Val Gln Leu Leu Glu Arg Phe Tyr Asp  
 1075 1080 1085  
 Pro Leu Ala Gly Lys Val Leu Leu Asp Gly Lys Glu Ile Lys Arg Leu  
 1090 1095 1100  
 Asn Val Gln Trp Leu Arg Ala His Leu Gly Ile Val Ser Gln Glu Pro  
 1105 1110 1115 1120  
 Ile Leu Phe Asp Cys Ser Ile Ala Glu Asn Ile Ala Tyr Gly Asp Asn  
 1125 1130 1135  
 Ser Arg Val Val Ser Gln Glu Glu Ile Val Arg Ala Ala Lys Glu Ala  
 1140 1145 1150  
 Asn Ile His Ala Phe Ile Glu Ser Leu Pro Asn Lys Tyr Ser Thr Lys  
 1155 1160 1165  
 Val Gly Asp Lys Gly Thr Gln Leu Ser Gly Gly Gln Lys Gln Arg Ile  
 1170 1175 1180  
 Ala Ile Ala Arg Ala Leu Val Arg Gln Pro His Ile Leu Leu Leu Asp  
 1185 1190 1195 1200  
 Glu Ala Thr Ser Ala Leu Asp Thr Glu Ser Glu Lys Val Val Gln Glu  
 1205 1210 1215  
 Ala Leu Asp Lys Ala Arg Glu Gly Arg Thr Cys Ile Val Ile Ala His  
 1220 1225 1230  
 Asn Leu Ser Thr Ile Gln Asn Ala Asp Ile Val Val Phe Gln Asn  
 1235 1240 1245  
 Lys Arg Val Lys Ile His Ile Thr His Gln Val Leu Leu Ala Gln Lys  
 1250 1255 1260  
 Gly Ile Tyr Phe Ser Met Val Ser Val Gln Ala Gly Thr Lys Arg Gln  
 1265 1270 1275 His.

4216 - 4  
 4217 - 1175  
 4218 - 1177  
 4219 - May 1990

Thr Gly Ile Gly Ala Gly Val Leu Ile Val Ala Tyr Ile Gln Val Ser  
                  20                 25                 30  
 Leu Thr Cys Leu Ala Ala Gly Arg Gln Ile His Lys Ile Arg Gln Lys  
                  35                 40                 45  
 Phe Phe His Ala Ile Met Asn Gln Glu Ile Gly Trp Phe Asp Val His  
                  50                 55                 60  
 Asp Val Gly Glu Leu Asn Thr Arg Leu Thr Asp Asp Val Ser Lys Ile  
                  65                 70                 75                 80  
 Asn Asp Gly Ile Gly Asp Lys Ile Gly Met Phe Phe Gln Ser Ile Thr  
                  85                 90                 95  
 Thr Phe Leu Ala Gly Phe Ile Ile Gly Phe Ile Ser Gly Trp Lys Leu  
                  100                105                110  
 Thr Leu Val Ile Leu Ala Val Ser Pro Leu Ile Gly Leu Ser Ser Ala  
                  115                120                125  
 Leu Trp Ala Lys Val Leu Thr Ser Phe Thr Asn Lys Glu Leu Gln Ala  
                  130                135                140  
 Tyr Ala Lys Ala Gly Ala Val Ala Glu Glu Val Leu Ala Ala Ile Arg  
                  145                150                155                160  
 Thr Val Ile Ala Phe Gly Gly Gln Lys Glu Leu Glu Arg Tyr Asn  
                  165                170                175  
 Lys Asn Leu Glu Glu Ala Lys Asn Val Gly Ile Lys Lys Ala Ile Thr  
                  180                185                190  
 Ala Ser Ile Ser Ile Gly Ile Ala Tyr Leu Leu Val Tyr Ala Ser Tyr  
                  195                200                205  
 Ala Leu Ala Phe Trp Tyr Gly Thr Ser Leu Val Leu Ser Asn Glu Tyr  
                  210                215                220  
 Ser Ile Gly Glu Val Leu Thr Val Phe Phe Ser Ile Leu Leu Gly Thr  
                  225                230                235                240  
 Phe Ser Ile Gly His Leu Ala Pro Asn Ile Glu Ala Phe Ala Asn Ala  
                  245                250                255  
 Arg Gly Ala Ala Phe Glu Ile Phe Lys Ile Ile Asp Asn Glu Pro Ser  
                  260                265                270  
 Ile Asp Ser Phe Ser Thr Lys Gly Tyr Lys Pro Asp Ser Ile Met Gly  
                  275                280                285  
 Asn Leu Glu Phe Lys Asn Val His Phe Asn Tyr Pro Ser Arg Ser Glu  
                  290                295                300  
 Val Gin Ile Leu Lys Gly Leu Asn Leu Lys Val Lys Ser Gly Gln Thr  
                  305                310                315                320  
 Val Ala Ile Val Gly Asn Ser Gly Cys Gly Lys Ser Thr Thr Val Gln  
                  330                335                340                345  
 Leu Met Ile Arg Leu Tyr Asp Ile Leu Ile Ile Val Val Val Ser Ile Asp  
                  350                355                360                365  
 Gly Ile Asp Ile Arg Thr Ile Asn Val Arg Tyr Leu Arg Asn Ile Ile  
                  375                380                385  
 Gly Val Val Ser Gln Glu Pro Val Leu Phe Ala Thr Thr Ile Ala Glu  
                  390                395                400  
 Asn Ile Arg Tyr Gly Arg Glu Asp Val Thr Met Asp Glu Ile Glu Lys  
                  405                410                415  
 Asn Val Ile Ile Ala Asn Ala Ile Arg Asn Ile Met Lys Ile Ile His

465                    470                    475                    480  
Ile Val Ile Ala His Arg Leu Ser Thr Val Arg Asn Ala Asp Val Ile  
485                    490                    495  
Ala Gly Phe Asp Gly Gly Val Ile Val Glu Gln Gly Asn His Asp Glu  
500                    505                    510  
Leu Met Arg Glu Lys Gly Ile Tyr Phe Lys Leu Val Met Thr Gln Thr  
515                    520                    525  
Arg Gly Asn Glu Ile Glu Pro Gly Asn Asn Ala Tyr Gly Ser Gln Ser  
530                    535                    540  
Asp Thr Asp Ala Ser Glu Leu Thr Ser Gln Glu Ser Lys Ser Pro Leu  
545                    550                    555                    560  
Ile Arg Arg Ser Ile Tyr Arg Ser Val His Arg Lys Gln Asp Gln Glu  
565                    570                    575  
Arg Arg Leu Ser Met Lys Glu Ala Val Asp Glu Asp Val Pro Leu Val  
580                    585                    590  
Ser Phe Trp Arg Ile Leu Asn Leu Asn Leu Ser Glu Trp Pro Tyr Leu  
595                    600                    605  
Leu Val Gly Val Leu Cys Ala Val Ile Asn Gly Cys Ile Gln Pro Val  
610                    615                    620  
Phe Ala Ile Val Phe Ser Arg Ile Val Gly Val Phe Ser Arg Asp Asp  
625                    630                    635                    640  
Asp His Glu Thr Lys Arg Gln Asn Cys Asn Leu Phe Ser Leu Phe Phe  
645                    650                    655  
Leu Val Met Gly Leu Ile Ser Phe Val Thr Tyr Phe Phe Gln Gly Phe  
660                    665                    670  
Thr Phe Gly Lys Ala Gly Glu Ile Leu Thr Lys Arg Val Arg Tyr Met  
675                    680                    685  
Val Phe Lys Ser Met Leu Arg Gln Asp Ile Ser Trp Phe Asp Asp His  
690                    695                    700  
Lys Asn Ser Thr Gly Ser Leu Thr Thr Arg Leu Ala Ser Asp Ala Ser  
705                    710                    715                    720  
Ser Val Lys Gly Ala Met Gly Ala Arg Leu Ala Val Val Thr Gln Asn  
725                    730                    735  
Val Ala Asn Leu Gly Thr Gly Val Ile Leu Ser Leu Val Tyr Gly Trp  
740                    745                    750  
Gln Leu Thr Leu Leu Val Val Ile Ile Pro Leu Ile Val Leu Gly  
755                    760                    765  
Gly Ile Ile Glu Met Lys Leu Leu Ser Gly Gln Ala Leu Lys Asp Lys  
770                    775                    780  
Tyr Ile Leu Ile Ile Ser Tyr Ile Leu Ile Thr Ile Ala Ile Gln Asn  
785                    790                    795                    800  
Ile Arg Thr Ile Val Ser Leu Ile Arg Ile Leu Lys Phe Ile Thr Met  
805                    810                    815  
Tyr Ala Gln Ser Leu Gln Ile Phe Tyr Arg Asn Ala Met Lys Lys Ala  
820                    825                    830  
His Val Phe Gly Ile Thr Phe Ser Phe Thr Gln Ala Met Met Tyr Phe  
835                    840                    845  
Ser Tyr Ala Ala Gly Phe Arg Phe Gly Ala Tyr Leu Val Ala Gln Ile  
850                    855                    860  
Ile Ile Val Ile Ile Ser Ile Ile Ser Ile Ile Val Ile Ile Val Ile

Glu Gly Asn Val Lys Phe Asn Gly Val Gln Phe Asn Tyr Pro Thr Arg  
930 935 940  
Pro Asn Ile Pro Val Leu Gln Gly Leu Ser Leu Glu Val Lys Lys Gly  
945 950 955 960  
Gln Thr Leu Ala Leu Val Gly Ser Ser Gly Cys Gly Lys Ser Thr Val  
965 970 975  
Val Gln Leu Leu Glu Arg Phe Tyr Asp Pro Met Ala Gly Ser Val Phe  
980 985 990  
Leu Asp Gly Lys Glu Ile Lys Gln Leu Asn Val Gln Trp Leu Arg Ala  
995 1000 1005  
His Leu Gly Ile Val Ser Gln Glu Pro Ile Leu Phe Asp Cys Ser Ile  
1010 1015 1020  
Ala Glu Asn Ile Ala Tyr Gly Asp Asn Ser Arg Ala Val Ser His Glu  
1025 1030 1035 1040  
Glu Ile Val Arg Ala Ala Lys Glu Ala Asn Ile His Gln Phe Ile Asp  
1045 1050 1055  
Ser Leu Pro Asp Lys Tyr Asn Thr Arg Val Gly Asp Lys Gly Thr Gln  
1060 1065 1070  
Leu Ser Gly Gly Gln Lys Gln Arg Ile Ala Ile Ala Arg Ala Leu Val  
1075 1080 1085  
Arg Gln Pro His Ile Leu Leu Asp Glu Ala Thr Ser Ala Leu Asp  
1090 1095 1100  
Thr Glu Ser Glu Lys Val Val Gln Glu Ala Leu Asp Lys Ala Arg Glu  
1105 1110 1115 1120  
Cly Arg Thr Cys Ile Val Ile Ala His Arg Leu Ser Thr Ile Gln Asn  
1125 1130 1135  
Ala Asp Leu Ile Val Val Ile Glu Asn Gly Lys Val Lys Glu His Gly  
1140 1145 1150  
Thr His Gln Gln Leu Leu Ala Gln Lys Gly Ile Tyr Phe Ser Met Val  
1155 1160 1165  
Gln Ala Gly Ala Lys Arg Ser  
1170 1175

<210> 5  
<211> 1066  
<212> PRT  
<213> Homo sapiens

- P.R.  
- 110 - PDB ID: 1D14  
- 30 - 1066-07-15

- 400 - 5

Trp Lys Leu Thr Leu Val Ile Met Ala Ile Ser Ile Ile Leu Gly Leu  
1 5 10 15  
Ser Ala Ala Val Trp Ala Lys Ile Leu Ser Ala Phe Ser Asp Lys Glu  
20 25 30  
Leu Ala Ala Tyr Ala Lys Ala Ile Ala Val Ala Glu Glu Ala Leu Gly  
35 40 45  
- 400 - 6

Lys Glu Tyr Thr Ile Gly Asn Ala Met Thr Val Phe Phe Ser Ile Leu  
115 120 125  
Ile Gly Ala Phe Ser Val Gly Gln Ala Ala Pro Cys Ile Asp Ala Phe  
130 135 140  
Ala Asn Ala Arg Gly Ala Ala Tyr Val Ile Phe Asp Ile Ile Asp Asn  
145 150 155 160  
Asn Pro Lys Ile Asp Ser Phe Ser Glu Arg Gly His Lys Pro Asp Ser  
165 170 175  
Ile Lys Gly Asn Leu Glu Phe Asn Asp Val His Phe Ser Tyr Pro Ser  
180 185 190  
Arg Ala Asn Val Lys Ile Leu Lys Gly Leu Asn Leu Lys Val Gin Ser  
195 200 205  
Gly Gin Thr Val Ala Leu Val Gly Ser Ser Gly Cys Gly Lys Ser Thr  
210 215 220  
Thr Val Gln Leu Ile Gln Arg Leu Tyr Asp Pro Asp Glu Gly Thr Ile  
225 230 235 240  
Asn Ile Asp GLY Gln Asp Ile Arg Asn Phe Asn Val Asn Tyr Leu Arg  
245 250 255  
Glu Ile Ile Gly Val Val Ser Gln Glu Pro Val Leu Phe Ser Thr Thr  
260 265 270  
Ile Ala Glu Asn Ile Cys Tyr Gly Arg Gly Asn Val Thr Met Asp Glu  
275 280 285  
Ile Lys Lys Ala Val Lys Glu Ala Asn Ala Tyr Glu Phe Ile Met Lys  
290 295 300  
Leu Pro Gln Lys Phe Asp Thr Leu Val Gly Glu Arg Gly Ala Gln Leu  
305 310 315 320  
Ser Gly Gly Gln Lys Gln Arg Ile Ala Ile Ala Arg Ala Leu Val Arg  
325 330 335  
Asn Pro Lys Ile Leu Leu Leu Asp Glu Ala Thr Ser Ala Leu Asp Thr  
340 345 350  
Glu Ser Glu Ala Glu Val Gln Ala Ala Leu Asp Lys Ala Arg Glu Gly  
355 360 365  
Arg Thr Thr Ile Val Ile Ala His Arg Leu Ser Thr Val Arg Asn Ala  
370 375 380  
Asp Val Ile Ala Gly Phe Glu Asp Gly Val Ile Val Glu Gln Gly Ser  
385 390 395 400  
His Ser Glu Leu Met Lys Lys Glu Gly Val Tyr Phe Lys Leu Val Asn  
405 410 415  
Met Asn Thr Ser Ile Asp Ile Ile Val Asp Val Ile Val Ile Asp  
41 41 41 41  
Asp Ile Lys Ala Ala Thr Arg Met Ala Phe Asn Val Trp Asp Ser Ala  
415 420 425 430  
Ile Phe Asp His Ser Thr Ile Ile Ile Asp Ile Lys Asp Ser Glu Met Gly  
440 445 450 455  
Gln Lys Ser Leu Asp Val Glu Val Asp Gly Leu Glu Ala Asn Val Pro  
465 470 475 480  
Pro Val Ser Phe Leu Lys Val Leu Lys Leu Asn Lys Thr Glu Trp Pro  
485 490 495 500  
Tyr Asp Val Val Asp Ile Ile Val Tyr Ala Ile Ala Asn Gly Leu Ile

565	570	575
Met Ala Phe Lys Ala Met Leu Arg Gln Asp Met Ser Trp Phe Asp Asp		
580	585	590
His Lys Asn Ser Thr Gly Ala Leu Ser Thr Arg Leu Ala Thr Asp Ala		
595	600	605
Ala Gln Val Gln Gly Ala Thr Gly Thr Arg Leu Ala Leu Ile Ala Gln		
610	615	620
Asn Ile Ala Asn Leu Gly Thr Gly Ile Ile Ser Phe Ile Tyr Gly		
625	630	635
Trp Gln Leu Thr Leu Leu Leu Ala Val Val Pro Ile Ala Val		
645	650	655
Ser Gly Ile Val Glu Met Lys Leu Leu Ala Gly Asn Ala Lys Arg Asp		
660	665	670
Lys Lys Glu Leu Glu Ala Ala Gly Lys Ile Ala Thr Glu Ala Ile Glu		
675	680	685
Asn Ile Arg Thr Val Val Ser Leu Thr Gln Glu Arg Lys Phe Glu Ser		
690	695	700
Met Tyr Val Glu Lys Leu Tyr Gly Pro Tyr Arg Asn Ser Val Gln Lys		
705	710	715
Ala His Ile Tyr Gly Ile Thr Phe Ser Ile Ser Gln Ala Phe Met Tyr		
725	730	735
Phe Ser Tyr Ala Gly Cys Phe Arg Phe Gly Ala Tyr Leu Ile Val Asn		
740	745	750
Gly His Met Arg Phe Arg Asp Val Ile Leu Val Phe Ser Ala Ile Val		
755	760	765
Phe Gly Ala Val Ala Leu Gly His Ala Ser Ser Phe Ala Pro Asp Tyr		
770	775	780
Ala Lys Ala Lys Leu Ser Ala Ala His Leu Phe Met Leu Phe Glu Arg		
785	790	795
Gln Pro Leu Ile Asp Ser Tyr Ser Glu Glu Gly Leu Lys Pro Asp Lys		
805	810	815
Phe Glu Gly Asn Ile Thr Phe Asn Glu Val Val Phe Asn Tyr Pro Thr		
820	825	830
Arg Ala Asn Val Pro Val Leu Gln Gly Leu Ser Leu Glu Val Lys Lys		
835	840	845
Gly Gln Thr Leu Ala Leu Val Gly Ser Ser Gly Cys Gly Lys Ser Thr		
850	855	860
Val Val Gin Leu Leu Glu Arg Phe Tyr Asp Pro Leu Ala Gly Thr Val		
865	870	875
Ile Leu Asp Ile Ile Ile Ala Tyr Ile Leu Asn Val Ile Trp Leu Asn		
885	890	895
Ala Phe Leu Ile Ile Val Ser Ile Leu Ile Ile Leu Ile Asp Ile Val		
905	910	915
Ile Ala Glu Asn Ile Ala Tyr Gly Asp Asn Ser Arg Val Val Ser Glu		
915	920	925
Asp Glu Ile Val Ser Ala Ala Lys Ala Ala Asn Ile His Pro Phe Ile		
930	935	940
Ile Thr Leu Phe His Ile Tyr Gln Thr Asn Val Ile Asp Ile Gly Thr		
945	950	955
Leu Asn Ile Ile Asn Asn Ile Ile Asn Ile Ile Asn Ile Ile Asn Ile		

Asn Ala Asp Leu Ile Val Val Phe Gln Asn Gly Arg Val Lys Glu His  
1025 1030 1035 1040  
Gly Thr His Gln Gln Leu Leu Ala Gln Lys Gly Ile Tyr Phe Ser Met  
1045 1050 1055  
Val Ser Val Gln Ala Gly Thr Gln Asn Leu  
1060 1065

<210> 6  
<211> 1266  
<212> PPT  
<213> Mus musculus

<300>  
<308> Genbank P21440  
<309> 1997-11-01

<400> 6  
Trp Lys Leu Thr Leu Val Ile Met Ala Ile Ser Pro Ile Leu Gly Leu  
1 5 10 15  
Ser Thr Ala Val Trp Ala Lys Ile Leu Ser Thr Phe Ser Asp Lys Glu  
20 25 30  
Leu Ala Ala Tyr Ala Lys Ala Gly Ala Val Ala Glu Glu Ala Pro Gly  
35 40 45  
Ala Ile Arg Thr Val Ile Ala Phe Gly Gly Gln Asn Lys Glu Leu Glu  
50 55 60  
Arg Tyr Gln Lys His Leu Glu Asn Ala Lys Lys Ile Gly Ile Lys Lys  
65 70 75 80  
Ala Ile Ser Ala Asn Ile Ser Met Gly Ile Ala Phe Leu Leu Ile Tyr  
85 90 95  
Ala Ser Tyr Ala Ala Phe Trp Tyr Gly Ser Thr Leu Val Ile Ser  
100 105 110  
Lys Glu Tyr Thr Ile Gly Asn Ala Met Thr Val Phe Phe Ser Ile Leu  
115 120 125  
Ile Gly Ala Phe Ser Val Gly Gln Ala Ala Pro Cys Ile Asp Ala Phe  
130 135 140  
Ala Asn Ala Arg Gly Ala Ala Tyr Val Ile Phe Asp Ile Ile Asp Asn  
145 150 155 160  
Asn Pro Lys Ile Asp Ser Phe Ser Glu Arg Gly His Lys Pro Asp Asn  
165 170 175  
Ile Ile Ile Asn Ile Ile Phe Ser Asp Val His Ile Phe Ser Tyr Phe Ser  
180 185 190  
Asp Ile Asn Ile  
195 200 205  
Gly Gln Thr Val Ala Leu Val Gly Asn Ser Gly Cys Gly Lys Ser Ile  
210 215 220  
Thr Val Gln Leu Leu Gln Arg Leu Tyr Asp Pro Thr Glu Gly Lys Ile  
225 230 235 240  
Ser Ile Asp Gly Gln Asp Ile Asn Asn Phe Asn Val Asp Cys Leu Arg  
245 250 255 260  
Ile Ile

Ser Gly Gly Gin Lys Gln Arg Ile Ala Ile Ala Arg Ala Leu Val Arg  
 325 330 335  
 Asn Pro Lys Ile Leu Leu Leu Asp Glu Ala Thr Ser Ala Leu Asp Thr  
 340 345 350  
 Glu Ser Glu Ala Glu Val Gln Ala Ala Leu Asp Lys Ala Arg Glu Gly  
 355 360 365  
 Arg Thr Thr Ile Val Ile Ala His Arg Leu Ser Thr Ile Arg Asn Ala  
 370 375 380  
 Asp Val Ile Ala Gly Phe Glu Asp Gly Val Ile Val Glu Gln Gly Ser  
 385 390 395 400  
 His Ser Glu Leu Met Lys Lys Glu Gly Ile Tyr Phe Arg Leu Val Asn  
 405 410 415  
 Met Gln Thr Ala Gly Ser Gln Ile Leu Ser Glu Glu Phe Glu Ala Arg  
 420 425 430  
 Ala Leu Val Arg Asn Pro Lys Ile Leu Leu Leu Asp Glu Ala Thr Ser  
 435 440 445  
 Ala Leu Asp Thr Glu Ser Glu Ala Val Val Gln Val Ala Leu Asp Lys  
 450 455 460  
 Ala Arg Lys Gly Arg Thr Thr Ile Val Ile Ala His Arg Leu Ser Thr  
 465 470 475 480  
 Val Arg Asn Ala Asp Val Ile Ala Gly Phe Asp Asp Gly Val Ile Val  
 485 490 495  
 Glu Lys Gly Asn His Asp Glu Leu Met Lys Glu Lys Gly Ile Tyr Phe  
 500 505 510  
 Lys Leu Val Thr Met Gln Thr Ala Gly Asn Glu Val Glu Leu Glu Asn  
 515 520 525  
 Ala Ala Ala Arg Ala Leu Val Arg Asn Pro Lys Ile Leu Leu Leu Asp  
 530 535 540  
 Glu Ala Thr Ser Ala Leu Asp Thr Glu Ser Glu Ala Val Val Gln Ala  
 545 550 555 560  
 Ala Leu Asp Lys Ala Arg Glu Gly Arg Thr Thr Ile Val Ile Ala His  
 565 570 575  
 Arg Leu Ser Thr Val Arg Asn Ala Asp Val Ile Ala Gly Phe Asp Gly  
 580 585 590  
 Gly Val Ile Val Glu Gln Gly Asn His Asp Glu Leu Met Arg Glu Lys  
 595 600 605  
 Gly Ile Tyr Phe Lys Leu Val Met Thr Gln Thr Arg Gly Asn Glu Ile  
 610 615 620  
 Ala Pro Gly Asn Asn Ala Val Gly Leu Ser Asp Glu Lys Ala Ala Gly  
 625 630 635 640  
 Arg Val Ala Phe Asn Gly Trp Lys Ala Arg Ile Phe Arg Asn Ser Thr  
 645 650 655 660  
 Lys Lys Ser Ile Lys Ser Ile His Gln Asn Asn Leu Asp Ala Glu Thr  
 660 665 670  
 Asn Glu Leu Asp Ala Asn Val Pro Pro Val Ser Phe Leu Lys Val Leu  
 675 680 685  
 Lys Leu Asn Lys Thr Glu Trp Pro Tyr Phe Val Val Gly Thr Val Cys  
 690 695 700  
 Ala Ile Ala Asn Gly Ala Ile Ile Ile Ala Ile Ser Ile Ile Ile Ile Ile

770	775	780
Gln Asp Met Ser Trp Phe Asp Asp His Lys Asn Ser Thr Gly Ala Leu		
785	790	795
Ser Thr Arg Leu Ala Thr Asp Ala Ala Gln Val Gln Gly Ala Thr Gly		800
805	810	815
Thr Lys Leu Ala Leu Ile Ala Gln Asn Thr Ala Asn Leu Gly Thr Gly		
820	825	830
Ile Ile Ile Ser Phe Ile Tyr Gly Trp Gln Leu Thr Leu Leu Leu		
835	840	845
Ser Val Val Pro Phe Ile Ala Val Ala Gly Ile Val Glu Met Lys Met		
850	855	860
Leu Ala Gly Asn Ala Lys Arg Asp Lys Lys Glu Met Glu Ala Ala Gly		
865	870	875
Lys Ile Ala Thr Glu Ala Ile Glu Asn Ile Arg Thr Val Val Ser Leu		880
885	890	895
Thr Gln Glu Arg Lys Phe Glu Ser Met Tyr Val Glu Lys Leu His Gly		
900	905	910
Pro Tyr Arg Asn Ser Val Arg Lys Ala His Ile Tyr Gly Ile Thr Phe		
915	920	925
Ser Ile Ser Gln Ala Phe Met Tyr Phe Ser Tyr Ala Gly Cys Phe Arg		
930	935	940
Phe Gly Ser Tyr Leu Ile Val Asn Gly His Met Arg Phe Lys Asp Val		
945	950	955
Ile Leu Val Phe Ser Ala Ile Val Leu Gly Ala Val Ala Leu Gly His		960
965	970	975
Ala Ser Ser Phe Ala Pro Asp Tyr Ala Lys Ala Lys Leu Ser Ala Ala		
980	985	990
Tyr Leu Phe Ser Leu Phe Glu Arg Gln Pro Leu Ile Asp Ser Tyr Ser		
995	1000	1005
Gly Glu Gly Leu Trp Pro Asp Lys Phe Glu Gly Ser Val Thr Phe Asn		
1010	1015	1020
Glu Val Val Phe Asn Tyr Pro Thr Arg Ala Asn Val Pro Val Leu Gln		
1025	1030	1035
Gly Leu Ser Leu Glu Val Lys Lys Gly Gln Thr Leu Ala Leu Val Gly		1040
1045	1050	1055
Ser Ser Gly Cys Gly Lys Ser Thr Val Val Gln Leu Leu Glu Arg Phe		
1060	1065	1070
Tyr Asp Pro Met Ala Gly Ser Val Leu Leu Asp Gly Gln Glu Ala Lys		
1075	1080	1085
Asp Ile Asn Ile Val Trp Leu Asp Ala Thr Leu Ile Ile Val Val Gln		
1090	1095	1100
Asp Ile Leu Phe Asp Tyr Ser Ile Ala Thr Asn Ile Asn Ile Ile		
1105	1110	1115
Asp Asn Ser Arg Val Val Pro His Asp Gln Ile Val Arg Ala Ala Lys		
1120	1125	1130
Glu Ala Asn Ile His Pro Phe Ile Gln Thr Leu Pro Gln Lys Tyr Asn		1135
1140	1145	1150
Thr Arg Val Gly Asp Lys Gly Thr Gln Ile Ser Gly Gln Lys Gln		
1155	1160	1165
Asp Asn Ser Arg Val Val Pro His Asp Gln Ile Val Arg Ala Ala Lys		

Glu Asn Gly Lys Val Lys Glu His Gly Thr His Gln Gln Leu Leu Ala  
1235 1240 1245  
Gln Lys Gly Ile Tyr Phe Ser Met Val Asn Ile Gln Ala Gly Thr Gln  
1250 1255 1260  
Asn Leu  
1265

<210> 7  
<211> 1207  
<212> PPT  
<213> Arabidopsis thaliana

<300>  
<305> Genbank A42150  
<309> 1997-03-13

<400> 7

Glu Lys Met Met Glu Glu Val Leu Lys Tyr Ala Leu Tyr Phe Leu Val  
1 5 10 15  
Val Gly Ala Ala Ile Trp Ala Ser Ser Trp Ala Glu Ile Ser Cys Trp  
20 25 30  
Met Trp Ser Gly Glu Arg Gln Thr Thr Lys Met Arg Ile Lys Tyr Leu  
35 40 45  
Glu Ala Ala Leu Asn Gln Asp Ile Gln Phe Phe Asp Thr Glu Val Arg  
50 55 60  
Thr Ser Asp Val Val Phe Ala Ile Asn Thr Asp Ala Val Met Val Gln  
65 70 75 80  
Asp Ala Ile Ser Glu Lys Leu Gly Asn Phe Ile His Tyr Met Ala Thr  
85 90 95  
Phe Val Ser Gly Phe Ile Val Gly Phe Thr Ala Val Trp Gln Leu Ala  
100 105 110  
Leu Val Thr Leu Ala Val Val Pro Leu Ile Ala Val Ile Gly Gly Ile  
115 120 125  
His Thr Thr Thr Leu Ser Lys Leu Ser Asn Lys Ser Gln Glu Ser Leu  
130 135 140  
Ser Gln Ala Gly Asn Ile Val Glu Gln Thr Val Val Gln Ile Arg Val  
145 150 155 160  
Val Met Ala Phe Val Gly Glu Ser Arg Ala Ser Gln Ala Tyr Ser Ser  
165 170 175  
Ala Ile Tyr Ile Ala Ile Ile Ile Tyr Ile Ile Tyr Ile Gly Ile Ala Ile  
185 190 195 200  
Ile Met Gly Ile Ile Ala Ile Tyr Ile Ile Ile Ile Ile Ile Ile Ile Ile  
195 200 205  
Leu Leu Leu Trp Tyr Gly Gly Tyr Leu Val Arg His His Leu Ile Asn  
210 215 220  
Gly Gly Leu Ala Ile Ala Thr Met Phe Ala Val Met Ile Gly Gly Leu  
225 230 235 240  
Ala Leu Gly Gln Ser Ala Pro Ser Met Ala Ala Phe Ala Ilys Ala Ilys  
245 250 255 260  
Ile Ile

Ala Leu Val Gly Ser Ser Gly Ser Gly Lys Ser Thr Val Val Ser Leu  
325 330 335  
Ile Gln Arg Phe Tyr Asp Pro Asn Ser Gly Gln Val Leu Leu Asp Gly  
340 345 350  
Gln Asp Leu Lys Thr Leu Lys Leu Arg Trp Leu Arg Gln Gln Ile Gly  
355 360 365  
Leu Val Ser Gln Glu Pro Ala Leu Phe Ala Thr Ser Ile Lys Glu Asn  
370 375 380  
Ile Leu Leu Gly Arg Pro Asp Ala Asp Gln Val Glu Ile Glu Glu Ala  
385 390 395 400  
Ala Arg Val Ala Asn Ala His Ser Phe Ile Ile Lys Leu Pro Asp Gly  
405 410 415  
Phe Asp Thr Gln Val Gly Glu Arg Gly Leu Gln Leu Ser Gly Gly Gln  
420 425 430  
Lys Gln Arg Ile Ala Ile Ala Arg Ala Met Leu Lys Asn Pro Ala Ile  
435 440 445  
Leu Leu Leu Asp Glu Ala Thr Ser Ala Leu Asp Ser Glu Ser Glu Lys  
450 455 460  
Leu Val Gln Glu Ala Leu Asp Arg Phe Met Ile Gly Arg Thr Thr Leu  
465 470 475 480  
Ile Ile Ala His Arg Leu Ser Thr Ile Arg Lys Ala Asp Leu Val Ala  
485 490 495  
Val Leu Gln Gln Gly Ser Val Ser Glu Ile Gly Thr His Asp Glu Leu  
500 505 510  
Phe Ser Lys Gly Glu Asn Gly Val Tyr Ala Lys Leu Ile Lys Met Gln  
515 520 525  
Glu Ala Ala His Glu Thr Ala Met Ser Asn Ala Arg Lys Ser Ser Ala  
530 535 540  
Arg Pro Ser Ser Ala Arg Asn Ser Val Ser Ser Pro Ile Met Thr Arg  
545 550 555 560  
Asn Ser Ser Tyr Gly Arg Ser Pro Tyr Ser Arg Arg Leu Ser Asp Phe  
565 570 575  
Ser Thr Ser Asp Phe Ser Leu Ser Ile Asp Ala Ser Ser Tyr Pro Asn  
580 585 590  
Tyr Arg Asn Gln Lys Leu Ala Phe Lys Asp Gln Ala Asn Ser Phe Trp  
595 600 605  
Arg Leu Ala Lys Met Asn Ser Pro Glu Trp Lys Tyr Ala Leu Leu Gly  
610 615 620  
Ser Val Gly Ser Val Ile Cys Gly Ser Leu Ser Ala Phe Phe Ala Tyr  
625 630 635 640  
Val Leu Ser Ala Val Ile Ser Val Tyr Tyr Asn Pro Asp His Glu Tyr  
645 650 655 660  
Met Ile Lys His Ile Asp Lys Tyr Cys Tyr Leu Leu Ile Gly Leu Ser  
660 665 670 675  
Ser Ala Ala Leu Val Phe Asn Thr Leu Gln His Ser Phe Trp Asp Ile  
675 680 685  
Val Gly Glu Asn Leu Thr Lys Arg Val Arg Glu Lys Met Leu Ser Ala  
690 695 700  
Val Leu Lys Asn Ile Met Ala Ile Asp Phe Asp Gln Ile Val Asn Ile Met  
705 710 715

770            775            780  
Lys Met Phe Met Thr Gly Phe Ser Gly Asp Leu Glu Ala Ala His Ala  
785            790            795            800  
Lys Gly Thr Gln Leu Ala Gly Glu Ala Ile Ala Asn Val Arg Thr Val  
805            810            815  
Ala Ala Phe Asn Ser Glu Ala Lys Ile Val Arg Leu Tyr Thr Ala Asn  
820            825            830  
Leu Glu Pro Pro Leu Lys Arg Cys Phe Trp Lys Gly Gln Ile Ala Gly  
835            840            845  
Ser Gly Tyr Gly Val Ala Gln Phe Cys Leu Tyr Ala Ser Tyr Ala Leu  
850            855            860  
Gly Leu Trp Tyr Ala Ser Trp Leu Val Lys His Gly Ile Ser Asp Phe  
865            870            875            880  
Ser Lys Thr Ile Arg Val Phe Met Val Leu Met Val Ser Ala Asn Gly  
885            890            895  
Ala Ala Glu Thr Leu Thr Leu Ala Pro Asp Phe Ile Lys Gly Gly Gln  
900            905            910  
Ala Met Arg Ser Val Phe Glu Leu Leu Asp Arg Lys Thr Glu Ile Glu  
915            920            925  
Pro Asp Asp Pro Asp Thr Thr Pro Val Pro Asp Arg Leu Arg Gly Glu  
930            935            940  
Val Glu Leu Lys His Ile Asp Phe Ser Tyr Pro Ser Arg Pro Asp Ile  
945            950            955            960  
Gln Ile Phe Arg Asp Leu Ser Leu Arg Ala Arg Ala Gly Lys Thr Leu  
965            970            975  
Ala Leu Val Gly Pro Ser Gly Cys Gly Lys Ser Ser Val Ile Ser Leu  
980            985            990  
Ile Gln Arg Phe Tyr Glu Pro Ser Ser Gly Arg Val Met Ile Asp Gly  
995            1000            1005  
Lys Asp Ile Arg Lys Tyr Asn Leu Lys Ala Ile Arg Lys His Ile Ala  
1010            1015            1020  
Ile Val Pro Gln Glu Pro Cys Leu Phe Gly Thr Thr Ile Tyr Glu Asn  
1025            1030            1035            1040  
Ile Ala Tyr Gly His Glu Cys Ala Thr Glu Ala Glu Ile Ile Gln Ala  
1045            1050            1055  
Ala Thr Leu Ala Ser Ala His Lys Phe Ile Ser Ala Leu Pro Glu Gly  
1060            1065            1070  
Tyr Lys Thr Tyr Val Gly Glu Arg Gly Val Gln Leu Ser Gly Gly Gln  
1075            1080            1085  
Ile Ile Arg Ile Ala Asn Ala Asn Ile Val Arg Lys Ala Ile Ile Ile  
1090            1095            1100  
Met Leu Ile Arg Ile Ala Ile Asn Ala Leu Asp Ala Ile Ser Ile Asn  
1105            1110            1115            1120  
Ser Val Gln Glu Ala Leu Asp Gln Ala Cys Ser Gly Arg Thr Ser Ile  
1125            1130            1135  
Val Val Ala His Arg Leu Ser Thr Ile Arg Asn Ala His Val Ile Ala  
1140            1145            1150  
Val Ile Asp Asp Gly Lys Val Ala Glu Ile Gly Ser His Ser His Leu  
1155            1160            1165  
Val Ile Asp Asp Gly Lys Val Ala Glu Ile Gly Ser His Ser His Leu

602 J. S. T. H. YU

<212> PRT

4213: *Arabidopsis thaliana*

<300>

<308> Genbank CAA71277

<309> 1997-05-19

<400> 8

Lys Cln Ala Ser His Arg Val Ala Lys Tyr Ser Leu Asp Phe Val Tyr  
 1 5 10 15  
 Leu Ser Val Ala Ile Leu Phe Ser Ser Trp Leu Glu Val Ala Cys Trp  
 20 25 30  
 Met His Thr Gly Glu Arg Gln Ala Ala Lys Met Arg Arg Ala Tyr Leu  
 35 40 45  
 Arg Ser Met Leu Ser Gln Asp Ile Ser Leu Phe Asp Thr Glu Ala Ser  
 50 55 60  
 Thr Gly Glu Val Ile Ser Ala Ile Thr Ser Asp Ile Leu Val Val Gln  
 65 70 75 80  
 Asp Ala Leu Ser Glu Lys Val Gly Asn Phe Leu His Tyr Ile Ser Arg  
 95 35 90 95  
 Phe Ile Ala Gly Phe Ala Ile Gly Phe Thr Ser Val Trp Gln Ile Ser  
 100 105 110  
 Leu Val Thr Leu Ser Ile Val Pro Leu Ile Ala Leu Ala Gly Gly Ile  
 115 120 125  
 Tyr Ala Phe Val Ala Ile Gly Leu Ile Ala Arg Val Arg Lys Ser Tyr  
 130 135 140  
 Ile Lys Ala Gly Glu Ile Ala Glu Glu Val Ile Gly Asn Val Arg Thr  
 145 150 155 160  
 Val Gln Ala Phe Thr Gly Glu Glu Arg Ala Val Arg Leu Tyr Arg Glu  
 165 170 175  
 Ala Leu Glu Asn Thr Tyr Lys Tyr Gly Arg Lys Ala Gly Leu Thr Lys  
 180 185 190  
 Gly Leu Gly Leu Gly Ser Met His Cys Val Leu Phe Leu Ser Trp Ala  
 195 200 205  
 Leu Leu Val Trp Phe Thr Ser Val Val Val His Lys Asp Ile Ala Asp  
 210 215 220  
 Gly Gly Lys Ser Phe Thr Thr Met Leu Asn Val Val Ile Ala Gly Leu  
 225 230 235 240  
 Ser Leu Gly Gln Ala Ala Phe Asp Ile Ser Ala Phe Val Arg Ala Lys  
 245 250 255  
 Ala Ala Ala Tyr Phe Ile Phe Lys Met Ile Val Asn Thr Val Thr  
 265 270 275 280  
 Lys Thr Ser Ala Lys Ser Gly Arg Lys Leu Glu Lys Val Asp Gly His  
 275 280 285  
 Ile Gln Phe Lys Asp Ala Thr Phe Ser Tyr Pro Ser Arg Pro Asp Val  
 290 295 300  
 Val Ile Phe Asp Arg Leu Asn Leu Ala Ile Phe Ala Ile Lys Ile Val  
 315 320 325 330

Ile Leu Tyr Gly Lys Asp Asp Ala Thr Ala Glu Glu Ile Thr Arg Ala  
385 390 395 400  
Ala Lys Leu Ser Glu Ala Ile Ser Phe Ile Asn Asn Leu Pro Glu Gly  
405 410 415  
Phe Glu Thr Gln Val Gly Glu Arg Gly Ile Gln Leu Ser Gly Gly Gln  
420 425 430  
Lys Gln Arg Ile Ala Ile Ser Arg Ala Ile Val Lys Asn Pro Ser Ile  
435 440 445  
Leu Leu Leu Asp Glu Ala Thr Ser Ala Leu Asp Ala Glu Ser Glu Lys  
450 455 460  
Ser Val Gln Glu Ala Leu Asp Arg Val Met Val Gly Arg Thr Thr Val  
465 470 475 480  
Val Val Ala His Arg Leu Ser Thr Val Arg Asn Ala Asp Ile Ile Ala  
485 490 495  
Val Val His Glu Gly Lys Ile Val Glu Phe Gly Asn His Glu Asn Leu  
500 505 510  
Ile Ser Asn Pro Asp Gly Ala Tyr Ser Ser Leu Leu Arg Leu Gln Glu  
515 520 525  
Thr Ala Ser Leu Gln Arg Asn Pro Ser Leu Asn Arg Thr Leu Ser Arg  
530 535 540  
Pro His Ser Ile Lys Tyr Ser Arg Glu Leu Ser Arg Thr Arg Ser Ser  
545 550 555 560  
Phe Cys Ser Glu Arg Glu Ser Val Thr Arg Pro Asp Gly Ala Asp Pro  
565 570 575  
Ser Lys Lys Val Lys Val Thr Val Gly Arg Leu Tyr Ser Met Ile Arg  
580 585 590  
Pro Asp Trp Met Tyr Gly Val Cys Gly Thr Ile Cys Ala Phe Ile Ala  
595 600 605  
Gly Ser Gln Met Pro Leu Phe Ala Leu Gly Val Ser Gln Ala Leu Val  
610 615 620  
Ser Tyr Tyr Ser Gly Trp Asp Glu Thr Gln Lys Glu Ile Lys Lys Ile  
625 630 635 640  
Ala Ile Leu Phe Cys Cys Ala Ser Val Ile Thr Leu Ile Val Tyr Thr  
645 650 655  
Ile Glu His Ile Cys Phe Gly Thr Met Gly Glu Arg Leu Thr Leu Arg  
660 665 670  
Val Arg Glu Asn Met Phe Arg Ala Ile Leu Lys Asn Glu Ile Gly Trp  
675 680 685  
Phe Asp Ala Val Asp Asn Thr Ser Ser Met Leu Ala Ser Arg Leu Glu  
690 695 700  
Asp Asn Ala Ile Leu Leu Lys Thr Ile Val Val Asp Arg Arg Val Thr Ile  
710 715 720  
Leu Leu Gln Asn Leu Val Val Thr Ser Phe Ile Ile Ala Phe  
725 730 735  
Ile Leu Asn Trp Arg Leu Thr Leu Val Val Leu Ala Thr Tyr Pro Leu  
740 745 750  
Val Ile Ser Gly His Ile Ser Glu Lys Leu Phe Met Gln Gly Tyr Gly  
755 760 765  
Ile Asp Leu Asn Leu Ala Ile Leu Lys Ala Asn Met Leu Ala Ile Val

835	840	845
Met Asp Lys Gly Leu Ala Gly Phe Lys Ser Val Met Lys Thr Phe Met		
850	855	860
Val Leu Ile Val Thr Ala Leu Ala Met Gly Glu Thr Leu Ala Leu Ala		
865	870	875
Pro Asp Leu Leu Lys Gly Asn Gln Met Val Ala Ser Val Phe Glu Ile		880
885	890	895
Leu Asp Arg Lys Thr Gln Ile Val Gly Glu Thr Ser Glu Glu Leu Asn		
900	905	910
Asn Val Glu Gly Thr Ile Glu Leu Lys Gly Val His Phe Ser Tyr Pro		
915	920	925
Ser Arg Pro Asp Val Val Ile Phe Arg Asp Phe Asp Leu Ile Val Arg		
930	935	940
Ala Gly Lys Ser Met Ala Leu Val Gly Gln Ser Gly Ser Gly Lys Ser		
945	950	955
Ser Val Ile Ser Leu Ile Leu Arg Phe Tyr Asp Pro Thr Ala Gly Lys		
965	970	975
Val Met Ile Glu Gly Lys Asp Ile Lys Leu Asp Leu Lys Ala Leu		
980	985	990
Arg Lys His Ile Gly Leu Val Gln Gln Pro Ala Leu Phe Ala Thr		
995	1000	1005
Thr Ile Tyr Glu Asn Ile Leu Tyr Gly Asn Glu Gly Ala Ser Gln Ser		
1010	1015	1020
Glu Val Val Glu Ser Ala Met Leu Ala Asn Ala His Ser Phe Ile Thr		
1025	1030	1035
Ser Leu Pro Glu Gly Tyr Ser Thr Lys Val Gly Glu Arg Gly Val Gln		1040
1045	1050	1055
Met Ser Gly Gln Arg Gln Arg Ile Ala Ile Ala Arg Ala Ile Leu		
1060	1065	1070
Lys Asn Pro Ala Ile Leu Leu Asp Glu Ala Thr Ser Ala Leu Asp		
1075	1080	1085
Val Glu Ser Glu Arg Val Val Gln Gln Ala Leu Asp Arg Leu Met Ala		
1090	1095	1100
Asn Arg Thr Thr Val Val Ala His Arg Leu Ser Thr Ile Lys Asn		
1105	1110	1115
Ala Asp Thr Ile Ser Val Leu His Gly Gly Lys Ile Val Glu Gln Gly		1120
1125	1130	1135
Ser His Arg Lys Leu Val Leu Asn Lys Ser Gly Pro Tyr Phe Lys Leu		
1140	1145	1150
Ter. Ser Ile Ser Gln Gln Ile Val Ile		
1155	1160	

• D110 - 9

• D111 - 986

• D120 - PPT

• D130 - Artificial Sequence

• D220 -

• M14 - Synthetic Intermediate sequence

Thr Asp Phe Ala Ser Lys Leu Glu Glu Glu Met Thr Arg Tyr Ala Tyr  
 50 55 60  
 Tyr Tyr Ser Gly Leu Gly Ala Gly Val Leu Val Ala Tyr Ile Gln Val  
 65 70 75 80  
 Ser Trp Leu Ala Ala Gly Arg Gln Ile Arg Lys Ile Arg Lys Phe Phe  
 85 90 95  
 His Ala Ile Leu Arg Gln Glu Ile Gly Trp Phe Asp Ile Thr Gly Glu  
 100 105 110  
 Leu Asn Thr Arg Leu Thr Asp Asp Ile Ser Lys Ile Asn Asp Gly Ile  
 115 120 125  
 Gly Asp Lys Val Gly Met Phe Phe Gln Val Ala Thr Phe Leu Ala Gly  
 130 135 140  
 Phe Ile Val Gly Phe Ile Gly Trp Lys Leu Thr Leu Val Ile Leu Ala  
 145 150 155 160  
 Ile Ser Pro Ile Ile Gly Leu Ser Ala Ala Val Trp Ala Lys Ile Leu  
 165 170 175  
 Ser Phe Ser Lys Glu Leu Ala Tyr Ala Lys Ala Gly Ala Val Ala Glu  
 180 185 190  
 Glu Leu Gly Ala Ile Arg Thr Val Ile Ala Phe Gly Gly Gln Lys Glu  
 195 200 205  
 Leu Glu Arg Tyr Gln Lys Leu Glu Ala Lys Lys Ile Gly Ile Lys Lys  
 210 215 220  
 Ala Ile Ser Ala Ile Ser Met Gly Ala Phe Leu Leu Ile Tyr Ala Ser  
 225 230 235 240  
 Tyr Ala Leu Ala Phe Trp Tyr Gly Ser Thr Leu Val Ile Ser Glu Tyr  
 245 250 255  
 Thr Ile Gly Ala Met Thr Val Phe Phe Ser Ile Leu Ile Gly Ala Phe  
 260 265 270  
 Ser Val Gly Gln Ala Ala Pro Ile Asp Ala Phe Ala Asn Ala Arg Gly  
 275 280 285  
 Ala Ala Tyr Ile Phe Lys Ile Ile Asp Asn Pro Ser Ile Asp Ser Phe  
 290 295 300  
 Ser Gly His Lys Pro Asp Ile Lys Gly Asn Leu Glu Phe Lys Asp Val  
 305 310 315 320  
 His Phe Ser Tyr Pro Ser Arg Glu Val Lys Ile Leu Lys Gly Leu Asn  
 325 330 335  
 Leu Lys Val Ser Gly Gln Thr Val Ala Leu Val Gly Ser Gly Cys Gly  
 340 345 350  
 Lys Ser Thr Thr Val Gln Leu Ile Gln Asn Leu Tyr Asp Pro Glu Gly  
 355 360 365  
 Val Ile Asp Gly Ile Asp Ile Asp Thr Ile Asn Val Asn Tyr Leu Asp  
 370 375 380  
 Ser Ile Ile Gly Val Val Ser Gly Gln Pro Val Leu Phe Ala Thr Thr  
 385 390 395 400  
 Ile Ala Glu Asn Ile Tyr Gly Arg Asp Val Thr Met Asp Glu Ile Glu  
 405 410 415  
 Lys Ala Val Lys Glu Ala Ala Asn Ala Tyr Glu Phe Ile Met Lys Leu Pro  
 420 425 430  
 Pro Asp Thr Leu Val Ile Gly Asn Ile Asp Ala Ile Leu Pro Thr Gly Ile  
 435 440 445

500	505	510
Glu Asp Gly Val Ile Val Glu Gly Ser His Asp Glu Leu Met Lys Lys		
515	520	525
Sly Val Tyr Phe Lys Leu Val Met Gln Thr Gly Ile Asn Ser Glu Ala		
530	535	540
Met Lys Ser Leu Arg Ser Ser Gln Asp Arg Asp Asp Leu Glu Val Pro		
545	550	555
Val Ser Phe Trp Arg Va. Leu Lys Leu Asn Thr Glu Trp Pro Tyr Val		
565	570	575
Val Gly Thr Val Cys Ala Ile Ile Asn Gly Leu Gln Pro Phe Ala Ile		
580	585	590
Ile Leu Ser Ile Ile Ala Val Phe Asp Asp Val Lys Leu Leu Leu		
595	600	605
Ala Val Val Pro Ile Ile Val Val Ala Gly Ile Val Glu Met Lys Leu		
610	615	620
Leu Gly Asn Ala Arg Asp Lys Lys Leu Glu Ala Gly Lys Ile Ala Thr		
625	630	635
Glu Ala Ile Glu Asn Ile Arg Thr Val Va. Ser Leu Thr Glu Lys Phe		
645	650	655
Glu Ser Met Tyr Leu Pro Tyr Arg Asn Ser Val Arg Lys Ala His Ile		
660	665	670
Tyr Gly Ile Thr Phe Ser Ile Ser Gln Ala Met Tyr Phe Ser Tyr Ala		
675	680	685
Gly Cys Phe Arg Phe Gly Ala Tyr Leu Val His Gly Leu Met Phe Val		
690	695	700
Ile Leu Val Phe Ser Ala Ile Val Leu Gly Ala Val Ala Leu Gly Thr		
705	710	715
Ser Ser Phe Ala Pro Asp Tyr Ala Lys Ala Lys Leu Ser Ala Ala Leu		
725	730	735
Phe Leu Ile Glu Arg Pro Ile Asp Ser Tyr Ser Glu Gly Leu Pro Asp		
740	745	750
Leu Glu Gly Val Phe Val Phe Asn Tyr Pro Thr Arg Pro Asp Val Pro		
755	760	765
Val Leu Gln Gly Leu Ser Leu Glu Val Lys Lys Gly Gln Thr Leu Ala		
770	775	780
Leu Val Gly Ser Ser Gly Cys Gly Lys Ser Thr Val Val Gln Leu Leu		
785	790	795
Glu Arg Phe Tyr Asp Pro Leu Ala Gly Lys Val Leu Leu Asp Gly Lys		
805	810	815
Ala Ile Lys Ile Leu Arg Val Ile Thr Leu Arg Ala His Leu Gly Ile		
820	825	830
Val Ser Gln Ile Pro Ile Val Ile Arg Tyr Ser Ile Ala Val Asn Ile		
835	840	845
Ala Tyr Gly Asp Asn Ser Arg Val Ser Asp Gln Ile Val Ala Ala Lys		
850	855	860
Ala Asn Ile His Phe Ile Glu Thr Leu Pro Asp Lys Tyr Thr Arg Val		
865	870	875
Tyr Asp Lys Gly Thr Gln Leu Ser Gly Gly Gln Lys Gln Arg Ile Ala		
885	890	895
Asn Ile Ile Asp Ile Ile Asn Ile		

Val Lys Glu His Gly Thr His Gln Gln Leu Leu Ala Gln Lys Gly Ile  
                   965                  970                  975  
 Tyr Phe Ser Met Val Val Gln Ala Gly Thr  
                   983                  986

```
<210> 10  
<211> 14000  
<212> DNA  
<213> Arabidopsis thaliana
```

```
<220:>
<221:> misc_feature
<222:> (3429)...(0)
<223:> Translation start codon
```

<303>  
<303> Genbank AP000386  
<309> 1999-08-03



ttcaaaagaga cgtgaatctt atgctaaccgc cgggtgttata gctgaggcagg taaaataaaaga 5700  
ttaaaggttt ttccctttatt tggttttattt aatgccccacc aactgttcga cgtaatgccc 5762  
tttaactgtt cttaggcttg cttttttgtt sttgatctt caaatgattt tttttttttt 5821  
ttatgttgc ttatgoatc taaaagactta aacagtgtttt taatataatc aaaatcttgc 5881  
cttttttttgc gtccggttaat gattgttgc ttgggttttgc agggaaatttgc tcaagtttgc 5940  
actgtttatt cttatgttgc agagagtaag gtaacttaatc cgtatttgc tgcgatttgc 6000  
tatacgctt aacgtcggtt taaaaggcccc atggcttaatc gggttgggtttt aggatgtact 6060  
tatggaaatag cttgtatgtt atgggttttgc gtgttttgc atgttggagt ttttatttgc 6120  
aatggacaaa cccatggagg aaaggccgtt actgtatata tcttgtat tttttgttgc 6181  
atgtaaat taaaaacata acaatgtatgtt ggttttgc attttgc tttttttttt 6241  
ttggatgtca ctttttgc tttgtgtgtt ttttaggat tttggggcaat ctttcttgc 6301  
tcttggggcg ttttagtaaag gtaaaaggcccc tgggttataatc ttgtatggaga taattaaacc 6361  
gagacggagg ataattttaaag acccgttggg tggaaaatgtt ttggatcaag ttcatggggaa 6420  
cattgagttt aaagatgtga ctttttagtta tttttccatgg cttgtatgttgc tttatgggg 6480  
gaacttcaat attttttttgc cttttggggaa aactgtgggg gttgttgggg 6541  
tggaaagagt actgttgc ttttttttttgc gagatttataatc gatccaaaca cccgttgc 6600  
gttttttttgc atttqagtgc ttggcttagtgc actggatcat aacttttttttgc 6661  
attttttttgc gatggcagg gcaaaatttgc ttggatgggttgc ttgagataaaa gacgttgc 6720  
ttgaagttt tgggtgaaca aatgggggttgc gtgaatcaag aacgttgc 6780  
actataactag agaacataact cttatggaaag cttgtatgc aatgttgc 6840  
gttgc ttttttttgc gcatatgttgc attacatttttgc ttcttaaagg cttacacaca 6900  
caggtataaa taaaaatattt gattatatgtt cttatgttgc tttatggcattt ttttttttttgc 6960  
ttttgttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 7020  
gaagcagaga attttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 7080  
tgaagtttaca aacgttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 7140  
agtttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 7200  
tggatccatt ggggttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 7260  
catttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 7320  
agattttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 7380  
aaccgttca ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 7440  
tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 7500  
tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 7560  
tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 7620  
tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 7680  
tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 7740  
tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 7800  
tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 7860  
tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 7920  
tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 7980  
tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 8040  
tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 8100  
tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 8160  
tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 8220  
tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 8280  
tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 8340  
tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 8400  
tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 8460  
tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 8520  
tttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 8580

gggatagagt	gttgcatatg	tgatttaact	atatggttct	acgtcatgtt	ttagtggcga	9120
cttagaccc	tgattgtcaa	tcttattttt	tacaagtgaa	ctattattac	catactgtgt	9180
tctaaatcat	aagtattaat	atatgtgggt	acattgcagg	aactatctct	gaagggtttt	9240
gttggagaca	caagttaaagg	tcatgtcaag	acttcataatg	ttgtctgtgt	aggaggcagt	9300
aaacattagaa	ccgttagcgg	tttcaatgca	cagagcaaga	tttcttcttt	gttctgtcat	9360
gagttcgtg	taaccctcaga	aaagaagctt	taaccgaagtc	aaaccttgggg	tttcttattt	9420
ggcctctcgc	agtttgcct	ctatggttt	gagggtttaa	tttctctggta	ttgtgcacccac	9480
tttgtgagta	aagggggtgc	aacccccc	aaagtgtatca	aagtgtttgt	gtttttggtc	9540
atctatgtcaa	actctgttgc	tgaaatgtt	agtttgcct	ctgtttttat	ttgggggggt	9600
gaagttgttg	gttgggttt	ctcggtttt	gacagggcaga	ccaggattga	ccggatgtat	9660
gttgcgtgtg	atccctgtgg	gacgtatcgt	ccggacattt	agtttaggc	tgttgcattt	9720
gtttaccctt	caagacccga	cgtcatgtt	tttccggact	ttaacccat	aattcgagct	9780
ggacatagcc	agatctttgt	gggggggggt	gggttcaagg	agatgtttgt	aattgcgtat	9840
atcgagccgt	tttacgaccc	tcttgcgtgg	aaagtgtatca	ttgtatggca	agacatccgc	9900
ccggctaaacc	tgaatcttct	aaaggctcaaa	atcggttttgc	ttcaacaaga	accagcttcc	9960
tttgcgtgtt	cgatcttgc	caacatggc	tatggtaaaag	atggtgcac	tgaatccgg	10020
gttaatgtat	cgatgtgggc	ccggaaatgt	caacgtttca	tcagtgtttt	acgttgcagg	10080
tataaaactc	cagtagggcga	aagaggagt	cgtttatcag	gtggacagaa	acagaggatc	10140
gcgtatgc	gagttgtgt	caagaaccc	acgttgcgtt	ttttagacga	agcaacttgc	10200
gcgtatgt	caaaatcaga	atgggtgtt	caagagggt	tagagagg	catgagagg	10260
ccggaccacc	ttgttagttc	tttccggat	tttccggat	tttccggat	tttccggat	10320
gtgattcaag	acggggccgt	ttgtggacaa	ggcagccatt	caagatgtgt	tagccgcacca	10380
gaggggagtt	atttcaaggt	ttttacatgt	ttttttatca	ggatttgcag	tttgcattat	10440
gattaaaaac	aaaaaatcgg	tttgcgttat	ttttttata	tttgcatttt	tttgcatttt	10500
atttctatgg	actttatcaga	ttatgtat	ttttttat	tttgcatttt	tttgcatttt	10560
ttatgtat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	10620
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	10680
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	10740
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	10800
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	10860
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	10920
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	10980
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	11040
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	11100
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	11160
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	11220
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	11280
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	11340
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	11400
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	11460
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	11520
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	11580
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	11640
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	11700
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	11760
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	11820
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	11880
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	11940
ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	ttttttat	12000

Sequence analysis of the DNA sequence above shows that it contains several potential restriction sites, including EcoRI, KpnI, PstI, SmaI, and XbaI. These sites are located at positions 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, and 10000 respectively.

tagtatgtgt atctgctttt tttgtatata ttaatgtaat agatagtgta gagaatcatg  
aggAACATTt tgaatcata gaactaaggc ataattttga tgggtgtaaa taataactatt  
atTTTGGTA asaaggatga ttggaccga attccgasta atTTTCCAGG tgaatatagt  
CATTAAGTAA CTTGATCACC ggattttaaa aatgaattat gaasatgata CCATTTAAAC  
taatcaaaaa aaaaatgaaaa aattcatgtg gaaatcatat CCTCAAATC TGTTAACC  
aaactacttt CATTGGTTt gtgtgagaaa tataagaatt atCAAACAAA gaaaaaggTT  
ttgtttaattt ataatgatac gaggttcata aaattttat accactaaaa tattttata  
CTCATCAAA tattttctc gcttacgtcag ttttaatata atgtacaaat cataccata  
AGGGPATTGG CGATGTAA aattttggaa tttagttgtt atttgtttt CTTGAGTCCT  
atgttcaaaa aatatatttg ataatctaata tataaaaaat taccataat cacattgatc  
aatgtttaaa gttttgtatt agttgttgc tttttgtttt ttgggtgtg ttgaattttc  
ataaaitgaaa agtaaataacg ataatccatata ttcaaaacgt aaagggttta attagtatgc  
taaatggtaa aatataaaaca gtatattata aaaattacta aaatgggtttt gtaaaaaaaa  
tatgttattt gtattataat gaaattttaaa aattttaaat agaaatgtata attttgtcac  
aaagaggttt tgaggtgtta taattcatga agtaaattttt atttactgac gggtaagttg  
tgaaaaagtt tgaagactat ttttttgc tttcaagagaaaactactt tcttttttaa  
ttgtatqatq aqaaggcAAA agtgcagaca tttttttttt ttttccat tttcaacaat  
gtcaactcgTT gtattattca tatttttagca aactggttat atctatataatc atcaatcatt  
ttagaacatc atatccatca gtttttggac attgtcacat acgttagtat tgatgtacca  
gttaccctaa caggcttttgc catagtgtgg cagaacacgt gaggtgttat atatgcggat  
gaatttctatg ttctgcattt ttttaccatt catataaaaag tatttttttgc tttgtgtgc  
ggtttaagtc ctaatacaga ttttatttgc tttttttttt tataaaccac attggaaaccc  
aaactctaag agaataataag atcaatgtgt aattaataaa tttttagtgc attcgatcaa  
aacttagaaa gataacaaaa gtaagaacgaa gtatTTTAA gcgaataactc ttttagatatt  
ctttgtattt atcattatttgc acaaaccatt aaacatTTTt gtagacccta aaatccgatc  
tccaaaaaac aattttttttt

(210) 11

1212 19

### 21S rRNA

(213) *Arabidopsis thaliana*

400-11

cactgctcaa tqatctcggt ttctca

29

(210) 12

(21): 30

:211: DNA

## 2.1 Arabidopsis thaliana

- 12 -

It means a great deal to me.

2

215

• 11 • (2)

### 421B: DNA

<213> Artificial Sequence

卷之三

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 14

tttctccata ttgaccatca tactcattg

29